

Genome Search

Genome Search is another way for selecting and browsing genomes in IMG.

Genomes can be searched via implicit attributes using the **Quick Genome Search** box at the top right corner of every IMG page, as shown in Figure 1.

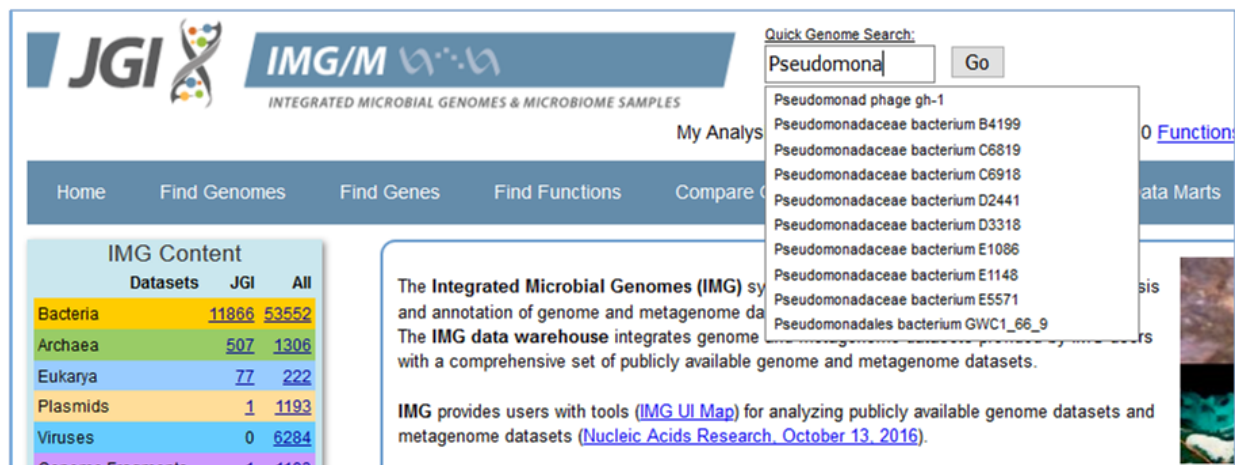


Figure 1. Quick Genome Search.

A detailed search of genomes can be carried out using **Genome Search** on the second-level menu of **Find Genomes**. There are 4 options: *by Fields*, *by Metadata Categories*, *by Metadata Category Operation*, and *by Metadata Category Chart*. The result of **Genome Search** is a list of genomes, which can be examined individually, or selected and saved for reducing the genome context for future analysis.

By Fields

Genome Search can use *genome fields* as search filter to find genomes, as shown in Figure 2.

Genome Search

by Fields | by Metadata Categories | by Metadata Category Operation | by Metadata Category Chart

Genome Search by Fields

Find genomes by keyword or substring.

Keyword

Filters

Genome Name / Sample Name

Genome Name / Sample Name

Study Name
NCBI Taxon ID (list)
RefSeq Project ID (list)
GenBank Project ID (list)
Scaffold External Accession (list)
Scaffold ID (list)
IMG Genome ID (IMG Taxon ID) (list)
IMG Submission ID (list)
JGI Project ID / ITS PID (list)
GOLD Study ID (list)
GOLD Sequencing Project ID (list)
GOLD Analysis Project ID (list)
Domain
Phylum
Sequencing Status
Sequencing Center
Funding Agency
IMG Release
All field names*

Examples

- "pseudomonas" as Genome Name retrieves all genomes with the substring "pseudomonas" such as "Pseudomonas syringae B728a".
- "62977" for NCBI Taxon ID retrieves "Acinetobacter sp. AD1".
- "NC_008009, NC_008010, NZ_AAKW0100001" as Scaffold External Accession will retrieve genomes that have scaffolds with those accession ID's.
- "arch" for Domain retrieves all Archaea.
- "proteo" for Phylum retrieves all proteobacteria.
- "finished" for Sequencing Status retrieves all finished genomes.
- "draft" for Sequencing Status retrieves all finished genomes.
- "JGI" for Sequencing Center retrieves all genomes sequenced at JGI.
- "2.0" for IMG Release retrieves all genomes with IMG release 2.0

hint:

The search
Use an underscore
Use % to match
All matches

word or part

Contact us
Accessibility/Section
Disclaimer

JGI is a DOE Office
© 1997-2017 The
JGI

Version 4.570 May 2017
gpweb37.nersc.gov/gemini2_shared/5.016000 2017-06-20-15:00:33 128.3.93.94

Berkeley National Laboratory

Figure 2. Genome Search by Fields.

Searchable fields are available in the pull-down filter list, including:

- Genome Name / Sample Name
- Study Name
- NCBI Taxon ID (list)
- RefSeq Project ID (list)
- GenBank Project ID (list)
- Scaffold External Accession (list)
- Scaffold ID (list)
- IMG Genome ID (IMG Taxon ID) (list)
- IMG Submission ID (list)
- JGI Project ID / ITS PID (list)
- GOLD Study ID (list)
- GOLD Sequencing Project ID (list)
- GOLD Analysis Project ID (list)
- Domain
- Phylum
- Sequencing Status
- Sequencing Center
- Funding Agency

- IMG Release
- all field names*

Filter options that include "(list)" allow users to enter multiple IDs separated by commas (,). All searches are case insensitive. Check "hint" for more guidance.

By Metadata Categories

Genome search by metadata categories allows users to search genomes using metadata fields from GOLD (see Figure 3). Searchable GOLD metadata categories/fields include:

- Altitude
- Biotic Relationships
- Cell Arrangement
- Cell Shape
- Culture Type
- Uncultured Type
- Depth
- Diseases
- Ecosystem
- Ecosystem Category
- Ecosystem Type
- Ecosystem Subtype
- Energy Source
- Geographic Location
- Gram Staining
- Habitat
- Host Gender
- Isolation
- Metabolism
- Motility
- Oxygen requirement
- Phenotype
- Relevance
- Salinity
- Specific Ecosystem
- Sporulation
- Temperature Range
- Type Strain
- PMO Project Id (list)
- ITS PID (list)
- ITS Proposal Id (list)

- GPTS Proposal Id (list)

Filter options that include "(list)" allow users to enter multiple IDs separated by commas (,). All searches are case insensitive. Check "hint" for more guidance.

Genome Search by Metadata Categories

Find genomes by metadata category keyword or substring.

Keyword

Filters

Altitude

Biotic Relationships

Cell Arrangement

Cell Shape

Culture Type

Uncultured Type

Depth

Diseases

Ecosystem

Ecosystem Category

Ecosystem Type

Ecosystem Subtype

Energy Source

Geographic Location

Gram Staining

hint:

The search is keyword based. Use an underscore to search for a word or part. Use % to match any characters. All matches are case insensitive.

Examples

- "singles" as Cell Arrangement retrieves all genomes with the "Singles".

- "ac" as Disease retrieves all genomes with the substring "ac" "Actinomyces", "Acne", "Lactococcus", etc.

[Contact us](#)

[Accessibility/Section 508](#)

[Disclaimer](#)

JGI is a DOE Office of Biological and Environmental Research

© 1997-2017 The DOE Office of Biological and Environmental Research

Version 4.570 Modified: 2017-05-16 10:25:43

[Berkeley National Laboratory](#)

7.56 67.180.254.83

Figure 3. Genome Search by Metadata Categories.

This search option is keyword based. If users do not have idea about what keywords to enter, then they should use the following two options (*metadata category operation* and *metadata category chart*) instead.

By Metadata Category Operation

This option allows users to search multiple conditions on the above mentioned metadata categories using "and" and/or "or" operators. Moreover, users can click on the plus-sign (+) to view pre-defined controlled vocabularies specified for each metadata category. For example, Figure 4 shows that **Biotic Relationships** can be: *Free living*, *Free-living*, *endophytic*, or *Symbiotic*. Users can click to select one or more controlled vocabularies. Conditions on multiple values of the same category will be combined using "or" operator.

It is also possible for users to specify conditions on multiple metadata categories (e.g., **Biotic Relationships** and **Cell Shape**). Conditions on different categories will be combined using "and" operator.

by Fields by Metadata Categories **by Metadata Category Operation** by Metadata Category Chart

hint: Expand tree to select or deselect values.

Select category search values.

Expand All Categories **Collapse All Categories**

(and: inter-category intersection; or: intra-category union)

Project Metadata	
<input type="checkbox"/> Altitude	or
and	
<input type="checkbox"/> Biotic Relationships <ul style="list-style-type: none"> <input type="checkbox"/> Free living <input type="checkbox"/> Free-living, endophytic <input type="checkbox"/> Symbiotic 	or
and	
<input type="checkbox"/> Cell Arrangement	or
and	
<input type="checkbox"/> Cell Shape	or
and	
<input type="checkbox"/> Culture Type	or

Figure 4. Genome Search by Metadata Category Operation.

It is possible that no genomes will be returned by a metadata category operation search because not all metadata category values have associated genomes in IMG, or there do not exist genomes in IMG satisfying combined search conditions.

To see which GOLD metadata category values actually have associated genomes in IMG, users can use the *metadata category chart* option described immediately below.

By Metadata Category Chart

To search genomes by metadata category chart, click the **by Metadata Category Chart** tab, select a category and then click the **Show Chart** button (see Figure 5(i)). Figure 5 (ii) shows the chart of **Cell Arrangement** and associated genome counts in IMG. Users can click on any bar to view the actual genome list. For example, Figure 5(iii) lists all genomes in IMG with *Tetrads* cell arrangement. Genomes can be selected to add to analysis cart or workspace for further analysis.

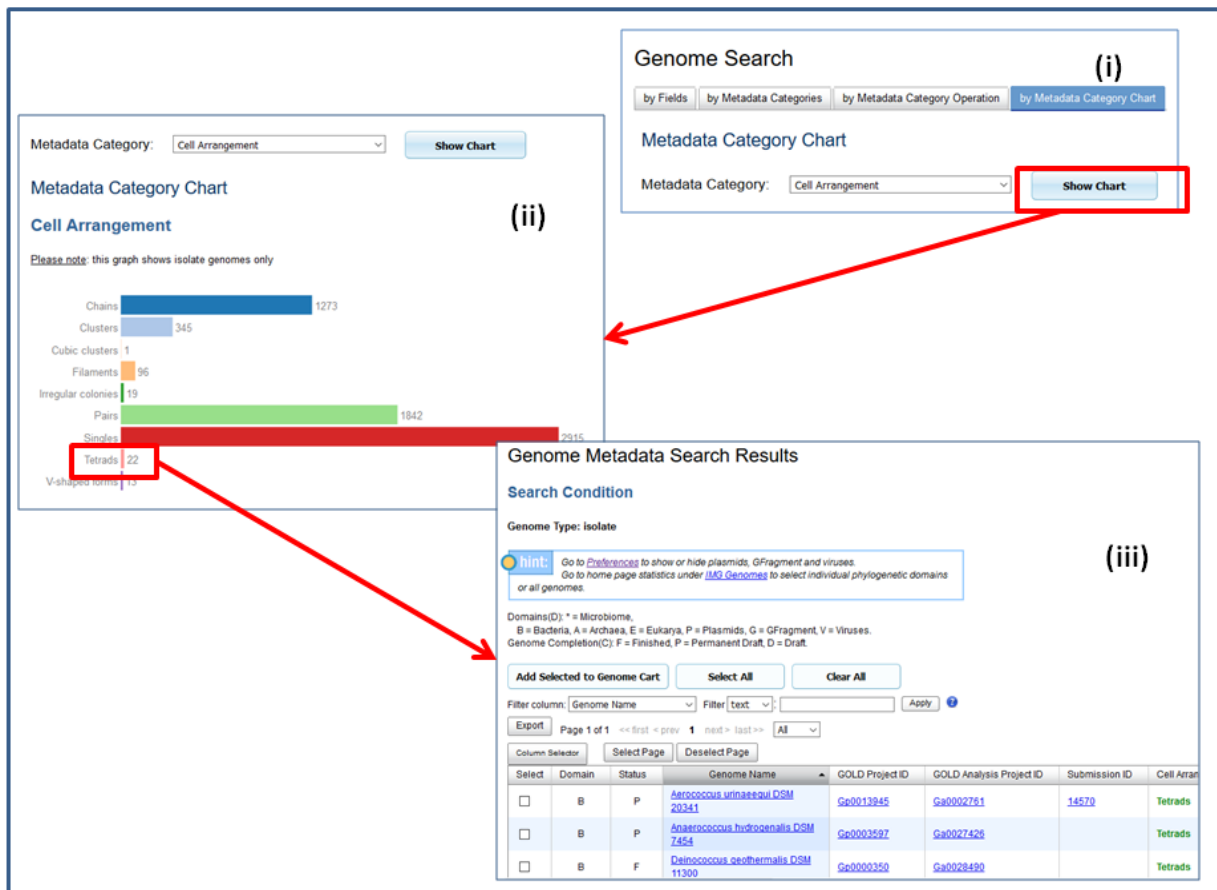


Figure 5. Genome Search by Metadata Category Chart.